



PAPER

The extraction of work from quantum coherence

Kamil Korzekwa^{1,3}, Matteo Lostaglio^{1,3}, Jonathan Oppenheim² and David Jennings¹¹ Department of Physics, Imperial College London, London SW7 2AZ, UK² Department of Physics, University College London, London WC1E 6BT, UK³ These authors contributed equally to this work.**Keywords:** quantum coherence, non-equilibrium thermodynamics, quantum thermodynamics, work extractionRECEIVED
22 July 2015REVISED
16 December 2015ACCEPTED FOR PUBLICATION
18 December 2015PUBLISHED
18 February 2016

Original content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](#).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



Abstract

The interplay between quantum-mechanical properties, such as coherence, and classical notions, such as energy, is a subtle topic at the forefront of quantum thermodynamics. The traditional Carnot argument limits the conversion of heat to work; here we critically assess the problem of converting coherence to work. Through a careful account of all resources involved in the thermodynamic transformations within a fully quantum-mechanical treatment, we show that there exist thermal machines extracting work from coherence arbitrarily well. Such machines only need to act on individual copies of a state and can be reused. On the other hand, we show that for any thermal machine with finite resources not all the coherence of a state can be extracted as work. However, even bounded thermal machines can be reused infinitely many times in the process of work extraction from coherence.

1. Introduction

Scientia potentia est, knowledge is power, the latin aphorism goes. This could not be more true in thermodynamics, where knowledge about the state of a system can be exploited to our advantage to extract work from it [1, 2]. In quantum mechanics states of maximal knowledge are called pure states. A peculiar feature of the quantum world is that, due to the superposition principle, even for such states there are many questions that cannot be answered sharply. In thermodynamics we are especially interested in energetic considerations and so an odd place is taken by pure states that are a superposition of different energy states. This is because, despite the fact that we possess full knowledge about the system, our possibility of predicting the outcome of an energy measurement can be very limited.

In standard quantum-mechanical considerations this is not a issue, because we can always reversibly transform a pure state into any other pure state by unitary dynamics. A basic task of thermodynamics, though, is the book-keeping of all energy flows from and out of the system, and there is no reversible transformation mapping a superposition of different energy states into an eigenstate while strictly conserving energy. Hence, we are left to wonder whether the ‘scientia’ of having a pure state with quantum coherence can be converted into ‘potentia’ of extracted work, while being limited by the law of energy conservation.

More precisely, we analyse work extraction from quantum coherence⁴, in the context of the theory of thermodynamics of individual quantum systems, currently under development [3–16]. The aim of the theory is to provide a suitable theoretical framework for our increasing ability to manipulate micro- and nanoscale systems [17–21]. Such general framework could also help approaching related questions, such as the role of quantum effects in biological systems [22–24] and the link between thermodynamics and quantum information processing. Although this field has recently seen a great number of contributions, most of the previously mentioned works do not incorporate the possibility of processing a state in a superposition of energy eigenstates. It was only relatively recently that the role of coherence in thermodynamics has been looked at more closely [25–34].

⁴ Here, and in the rest of the paper, we use the term ‘coherence’ in the sense of ‘superposition of states belonging to different energy eigenspaces’.

In this paper we first set the scene by presenting existing approaches to the problem of work extraction from the coherence of quantum systems. We argue that, within the regime of individually processed systems, the current approaches fail to account for all the resources used during the work extraction protocol. The typical assumption is to use a classical external field that experiences no back-reaction [35, 36]. However, this does not allow for a full accounting of the thermodynamic cost of maintaining the field. Although this cost may be small in a single use, it has to be accounted for since the work gain will also be small. Hence we propose an alternative framework that aims for a careful book-keeping of resources.

In particular, we use the notion of a *thermal machine* [37], a device of bounded resources that can be used to manipulate thermodynamical systems and perform tasks such as work extraction. Our thermal machine incorporates the use of an ancillary system carrying coherence (henceforth called *reference system*), introduced into the context of coherence manipulation in thermodynamics in [26]. It also includes a battery system where work can be stored, or transferred to the reference when necessary. A crucial question addressed in this work will be how to use the thermal machine in a repeatable way, i.e. without deteriorating it. We make use of an important result of Johan Åberg, showing that reference systems can be used repeatedly to manipulate coherence [31]. However the reference needs to be repumped for the machine to continue to operate. The work cost of repumping needs to be taken into account, which can mean that the thermal machine will not be able to extract all the available work from coherence. Nonetheless, we will find that we can come arbitrarily close, by choosing the amount of coherence resources carried by the reference system in the machine appropriately large.

On the other hand, for any given thermal machine, we will prove that one can never extract all the available work. We will show that coherences of individual quantum systems can be exploited to enhance the performance of work extraction protocols (both in the average and the single-shot sense), but not to the extent that could be expected in the ‘classical’ limit. Moreover, the work extraction protocol we provide does not deteriorate the thermal machine.

2. Coherence and work

2.1. Setting the scene

Let us start by introducing the framework that we will use throughout this paper and collect the core assumptions that our results rest upon. We want to study the allowed thermodynamic transformations by explicitly modelling any coherence resources being used. There are two ways in which coherence can enter the thermodynamics of the systems under consideration. This can happen either explicitly, by transferring it from an external system with quantum coherence (a trivial example being a swap operation between the system and an ancillary coherent state); or implicitly, by allowing operations that do not conserve energy (e.g., $|0\rangle \rightarrow |+\rangle$, where the Hamiltonian is given by $H_S = |1\rangle\langle 1|$) or conserve it only on average (e.g., $|1\rangle \rightarrow (|0\rangle + |2\rangle)/\sqrt{2}$, with $H_S = |1\rangle\langle 1| + 2|2\rangle\langle 2|$). Therefore, we will only allow for those transformations that do not implicitly introduce coherence:

Assumption 1. (Allowed transformations). The set of allowed transformations is given by all (strictly) energy-preserving unitaries, i.e., unitaries that commute with the total free Hamiltonian of the system. The use of all ancillary systems should be *explicitly* accounted for⁵.

We will also take a closer look at an alternative approach in section 2.4.1 and explain why we find it not satisfactory for the aims of the present work.

In this paper we focus on the task of work extraction from quantum systems with coherence. We do not aim here to settle the long-standing issue of what is an appropriate definition of work in quantum thermodynamics (see, e.g., [38, 39]). For the scope of this paper we will assume, for the sake of simplicity, that the following holds for classical (incoherent) states:

Assumption 2. (Average work, incoherent states). Let ρ_S be a quantum state of the system described by Hamiltonian H_S , with ρ_S being incoherent in the energy eigenbasis. Then, in the presence of a heat bath at temperature T , an average amount of work $\langle W \rangle(\rho_S)$ equal to the change of free energy of a state can be extracted from it:

$$\langle W \rangle(\rho_S) = \Delta F(\rho_S) := F(\rho_S) - F(\gamma_S), \quad (1)$$

where $F(\sigma) := \text{Tr}(\sigma H_S) - kTS(\sigma)$, $S(\cdot)$ is the von Neumann entropy and $\gamma_S = e^{-H_S/kT}/Z_S$ is a thermal state with Z_S being the partition function of the system.

⁵ As we will see, thermal ancillas are the only ones that can be freely introduced without trivialising the problem of work extraction.

This formula, consistent with traditional thermodynamics, has been obtained using work extraction models that differ in details, but agree on the result [9, 35]. For example, in [9] the work extraction protocol is based on two elementary processes: level transformations (that change the eigenvalues of the system Hamiltonian H_S) and full thermalisation with respect to the current system Hamiltonian (through thermal contact with a bath at temperature T). Average work is then defined as the average change in energy during level transformations (the ‘unitary’ steps) and, if initial and final Hamiltonian coincide, equation (1) is recovered. Here we focus on the problem of extending equation (1) to quantum states with coherence. The results of the present paper apply to any definition of work satisfying assumption 2.

The problem of work extraction can also be studied in the so-called single-shot regime. This means that one is interested in single instances of the work extraction protocol, instead of average quantities. To explain this more precisely, let us refer again to the model introduced in [9] that we have summarised above. Extracted work can then be seen as a random variable, maximising the average of which yields equation (1). However, we may instead ask what is the maximum amount of deterministic (i.e., fluctuation-free) work that can be extracted during a single instance of the protocol, while allowing the failure probability ϵ . In [9] it was shown that for incoherent states this quantity is given by $F_0^\epsilon(\rho_S) - F(\gamma_S)$, where $F_0^\epsilon(\rho_S) = -kT \log Z_\epsilon$ is a single-shot free energy defined as follows. Given a subset Λ of the indices $\{i\}$ labelling the energy levels of the system, define $Z(\Lambda) = \sum_{i \in \Lambda} e^{-\beta E_i}$, where E_i are the eigenvalues of H_S . Then $Z_\epsilon = \min_{\Lambda} \{Z(\Lambda) : \sum_{i \in \Lambda} p_i > 1 - \epsilon\}$, where $p_i = \langle E_i | \rho_S | E_i \rangle$. This result is also in agreement with other work extraction models based on thermal operations [7]. Hence, for the single-shot scenario we can use the following assumption:

Assumption 3. (Single-shot work, incoherent states). Let ρ_S be a quantum state of the system described by Hamiltonian H_S , with ρ_S being incoherent in the energy eigenbasis. Then in the presence of a heat bath at temperature T , using a single-shot protocol one can extract a sharp amount of work W_{ss}^ϵ with failure probability ϵ :

$$W_{ss}^\epsilon(\rho_S) = \Delta F_0^\epsilon(\rho_S) := F_0^\epsilon(\rho_S) - F(\gamma_S). \quad (2)$$

Once again, our aim is to extend equation (2) to quantum states with coherence and our results apply to any definition of single-shot work satisfying assumption 3. For the sake of brevity in the remaining of this paper we will only write ‘extracting work equal to the free energy’, omitting ‘in the presence of a heat bath at temperature T ’; however, this is how our claims should be understood.

In thermodynamic considerations thermal Gibbs states are the only ancillary states that can be introduced without the need for careful accounting. In fact, one can show that using energy-preserving unitaries (in accordance with assumption 1), a thermal state is the only one that can be introduced for free without allowing the production of every incoherent state [8]. Clearly, if this was possible, then from assumptions 2 and 3 one could extract infinite amount of work, thus trivialising the theory. Hence, the most general thermodynamic transformations that can be performed without using extra resources are given by:

- (i) adding a bath system in a thermal state γ_E with arbitrary Hamiltonian H_E and fixed temperature T ,

$$\rho_S \mapsto \rho_S \otimes \gamma_E, \quad \gamma_E = e^{-\beta H_E} / \text{Tr}(e^{-\beta H_E}), \quad \beta = 1/kT; \quad (3)$$

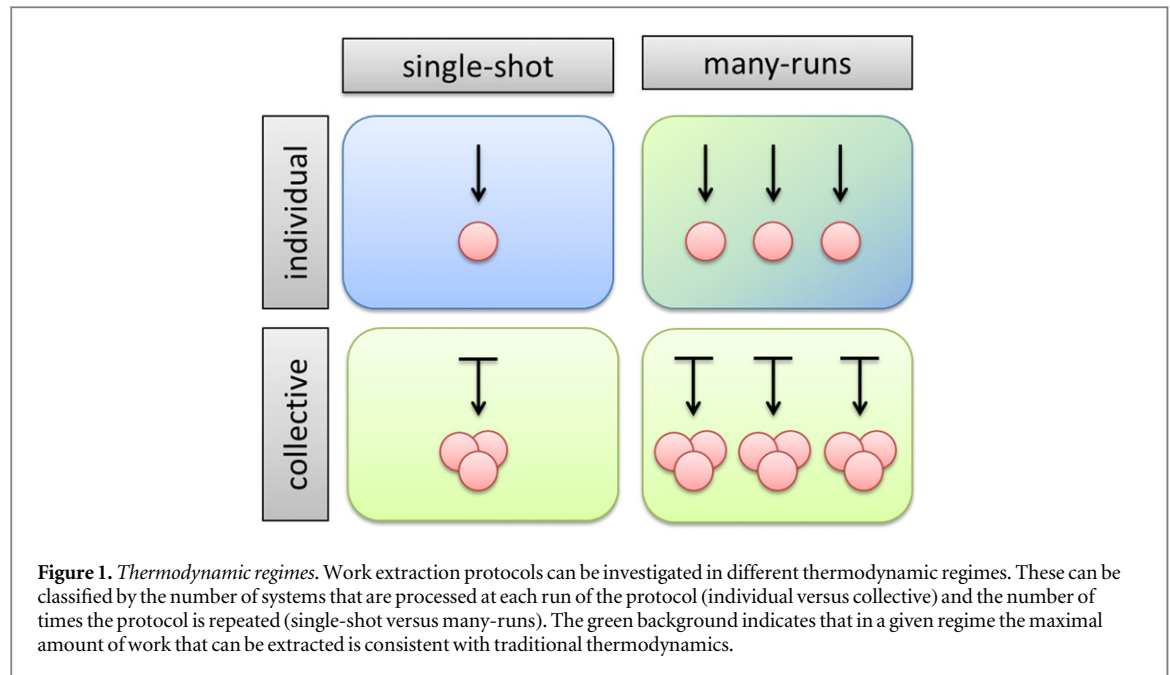
- (ii) performing any global unitary that conserves total energy, i.e., that commutes with the total free Hamiltonian of the system and baths, in accordance with assumption 1;
- (iii) discarding any subsystem.

The set of quantum maps acting on a system that arise from combining the transformations described above is known under the name of *thermal operations* [3, 26].

2.2. Work-locking

The aim of this work is to begin with a system initially in a state with coherence ρ_S , and finish with a thermal state γ_S , while optimally increasing the free energy of a battery (storage) system. The initial and final battery states, ρ_B and ρ'_B , should be incoherent, so that using assumptions 2 and 3 we can achieve the coherence to work conversion that we are looking for. Schematically:

$$\rho_S \otimes \rho_B \rightarrow \gamma_S \otimes \rho'_B. \quad (4)$$



Without the use of an ancillary resource state the above transformation is given by a thermal operation. Note that thermal operations commute with the dephasing channel \mathcal{D} [26] that removes all coherence from a quantum state

$$\mathcal{D}(\sigma) := \sum_i \text{Tr}(\Pi_i \sigma) \Pi_i,$$

where Π_i are the projectors on the energy eigenspaces of the system under consideration. Hence, we get that if the transformation described by equation (4) is possible, also the following one is:

$$\mathcal{D}(\rho_S) \otimes \rho_B \rightarrow \gamma_S \otimes \rho'_B.$$

This implies $\Delta F(\rho_B) \leq \Delta F(\mathcal{D}(\rho_S))$, because F is non-increasing under thermal operations. From assumption 2 we then have $\langle W \rangle(\rho_S) \leq \langle W \rangle(\mathcal{D}(\rho_S))$. A similar argument gives also $W_{ss}^\epsilon(\rho_S) \leq W_{ss}^\epsilon(\mathcal{D}(\rho_S))$; note that in both cases the bound is achievable because dephasing is a thermal operation. This phenomenon was observed before [7, 10] and was called ‘work-locking’ in [28]. Work-locking highlights that, despite contributing to the free energy of the state, quantum coherence does not contribute to work extraction: it is ‘locked’. It also shows, in agreement with [31], that the standard formula $\langle W \rangle(\rho_S) = \Delta F(\rho_S)$ applied to every state (also the ones with coherence), implicitly assumes the access to an external source of coherence. In this paper we revise the problem of extracting work from coherence, clarifying the role of this external source of coherence. To summarise

Central question. To what extent can equations (1) and (2) be extended to arbitrary quantum states with coherence, while explicitly accounting for coherence resources (ancillary systems)?

2.3. Different thermodynamic regimes

With an increasing interest in the thermodynamics of non-equilibrium quantum systems, an important distinction to make is between ‘single-shot’ statements, which are valid for every run of the protocol, and ‘many-runs’ statements, valid in the case of a large number M of runs. In the *asymptotic regime* $M \rightarrow \infty$ one is focused on studying average quantities (like average extracted work), which is justified by the fact that the fluctuations around the average can be made negligible in the limit of a large number of runs of the protocol (which is often the situation of interest in the study of heat engines). On the other hand, although the expected amount of extracted work can be studied in a single-shot regime [35], it potentially carries little information about the system at hand due to the large fluctuations of non-equilibrium thermodynamics [9]. Instead, the focus in the single-shot regime is typically on probabilistic work extraction protocols that guarantee precise and sharp amount of work with a finite probability of success or some minimum amount of guaranteed work [5, 7, 9, 11]. On top of this classification we can also differentiate between ‘individual processing’ scenarios, in which a single (possibly nanoscale) system undergoes a thermodynamic process on its own; and ‘collective’ scenarios, in which $N > 1$ copies of a state are processed together (the $N \rightarrow \infty$ limit is considered in [26]).

This classification of the thermodynamic regimes in which work extraction can be analysed is presented in figure 1. The work-locking described in the previous section is a feature appearing in the regime of individually

processed quantum systems. On the other hand, allowing for collective processing of the systems or for an ancillary quantum memory, one can ‘unlock’ work from quantum coherence (see appendix A for a short description of this point). For arbitrary collective processes, one only needs to use a sublinear amount of coherence in the reference system, meaning that its consumption does not contribute to the average work consumed or produced in the $N \rightarrow \infty$ limit [26]. In this limit coherence plays no role as, e.g., for N identical qubits $F(\rho^{\otimes N}) \approx F(\mathcal{D}(\rho^{\otimes N}))$, with a deficit scaling as $\log N/N$ [28]. Hence, in this paper we are interested in the thermodynamics of individual quantum systems ($N = 1$, the upper half of figure 1).

2.4. Individual processing regime

2.4.1. Average energy conservation

In [35] sharp energy conservation, as expressed by the unitary dynamics commuting with the total Hamiltonian H_{tot} , was replaced with the condition that such dynamics only keeps the first moment $\langle H_{\text{tot}} \rangle$ constant. Under this weaker condition it was shown that an amount of energy equal to the free energy difference $\Delta F(\rho_S)$ can be extracted on average from a system in an arbitrary quantum state ρ_S .

The elegance and appeal of this is that it recovers a clear thermodynamic meaning for the free energy of an individual quantum system. However, several problematic issues can be raised. Firstly, if one is interested in analyzing the class of allowed quantum operations, then in the average-energy scenario this set depends on the particular state one is processing, which is conceptually less appealing and technically problematic from a resource-theoretic perspective.

Secondly, restricting energy considerations to the first moment analysis can hide arbitrarily large energy fluctuations described by higher moments, that are not explicitly modeled, but may be highly relevant. To see this consider a unitary U_{ave} mapping a state $|\psi_{02}\rangle := (|0\rangle + |2\rangle)/\sqrt{2}$ to $|1\rangle$, which preserves energy on average (here $|n\rangle$ is the energy eigenstate corresponding to energy n). Since microscopically all processes are ultimately energy-conserving, U_{ave} must be realised through a joint energy-preserving unitary U involving $|\psi_{02}\rangle$ and some ancillary state ρ_A , e.g., the state of the battery

$$U(|\psi_{02}\rangle\langle\psi_{02}| \otimes \rho_A)U^\dagger = |1\rangle\langle 1| \otimes \rho'_A.$$

In any such process the energy fluctuations of the ancillary system must increase. Specifically, denoting by H the Shannon entropy of the outcomes of an energy measurement one gets a strict inequality: $H(\rho'_A) > H(\rho_A)$ (see appendix B for details). As an example consider the ancillary system prepared in the energy eigenstate $|m\rangle$, so that an energy measurement would give a sharp outcome. Then, while the system is transformed from $|\psi_{02}\rangle$ into $|1\rangle$, the ancilla must be transformed into a superposition of energy eigenstates $(|m+1\rangle + |m-1\rangle)/\sqrt{2}$. Hence an energy measurement would show fluctuations in the final state of the ancilla. It is important to note that the protocol that extracts work from coherence within the framework of average energy conservation [35], necessarily creates such extra fluctuations, however these are not explicitly modeled within the formalism used. As we will see it is exactly due to these fluctuations that our protocols require work to be invested in restoring the ancillary state.

Finally, as the fluctuations created by operations that conserve energy only on average remain outside the formalism, one cannot properly account for the fluctuations in the extracted work outside the asymptotic regime.

2.4.2. Repeatable use of coherence resources

As already mentioned in section 2.2, in the presence of energy conservation and without additional coherence resources, work-locking prevents us from extracting work from the coherence of individual quantum systems. One could then stay within the framework of strict energy conservation, but allow for the use of an extra source of coherence. We refer to this extra system as the *reference*.

At one extreme one could allow for the use of an infinite source of coherence (an ‘unbounded’ reference frame [40, 41])⁶, that entirely negates the constraints and experiences no back-reaction from its use on the quantum system. As suggested in [28], in such case we should be able to extract all the work from coherence. However, one might worry that this involves the accounting ‘ $\infty - c = \infty$ ’, with c being some finite resource consumed from an infinitely large reference system. Indeed, the use of such an unbounded reference allows us to simulate the operations from the previous section (conserving energy only on average) [41], and hide the arising extra fluctuations in the infinitely big reference system. This semiclassical treatment is typical for many standard approaches that assume the existence of a classical field experiencing no back-reaction from the system [35, 36] and works well in many circumstances. However we are interested in the regime in which the thermal machine itself may be a microscopic quantum system. Hence, it seems more reasonable to firstly consider the reference as a quantum system with finite coherence resources—a ‘bounded’ reference frame—and only then study the limit

⁶ Not to be confused with a reference described by a Hamiltonian unbounded from below, which is unphysical.

of an unbounded reference (recent works in this spirit and the discussion of semiclassical approaches can be found, e.g., in [29, 38, 42]).

Definition 1. (Reference). We consider a reference (or coherence reservoir) given by an infinite-dimensional ladder system described by Hamiltonian $H_R = \sum_{n=0}^{\infty} n |n\rangle\langle n|$. We characterise the state ρ_R of the reference through two numbers, $(\langle\bar{\Delta}\rangle, M)$. The first parameter, $\langle\bar{\Delta}\rangle$, measures the coherence properties of the reference and is given by

$$\langle\bar{\Delta}\rangle = \text{Tr}(\rho_R \bar{\Delta}), \quad \bar{\Delta} = (\Delta + \Delta^\dagger)/2, \quad (5)$$

where Δ is the shift operator $\Delta = \sum_{n=0}^{\infty} |n+1\rangle\langle n|$. We have that $\langle\bar{\Delta}\rangle < 1$ and the limit case $\langle\bar{\Delta}\rangle = 1$ is called *unbounded* or classical reference. The second parameter, M , describes the lowest occupied energy state, $M = \min\{n : \langle n|\rho_R|n\rangle > 0\}$.

Examples of a sequence of references that come arbitrarily close to a classical one are uniform superpositions of L energy states when $L \rightarrow \infty$ or coherent states with arbitrarily large amplitude. The use of $\langle\bar{\Delta}\rangle$ and M as relevant quality parameters will soon become clear.

Results from the field of quantum reference frames [40, 41, 43–45] suggest that the back-reaction experienced by the reference will necessarily deteriorate it and consume the resources. However, if the usefulness of the reference or field is continually degraded during the work extraction process, we cannot claim that we are presenting a protocol performing work extraction from the state alone, as extra resources are consumed. Similar problems arise if free energy is continually taken away from the reference.

In this paper we propose the following approach. We allow for the use of additional coherence resources as part of our thermal machine, but demand that they are used *repeatably* in the following sense: the performance of our reference-assisted protocol, while operating individually on the n th copy of the system, must be the same as while operating on the $(n+1)$ th copy, for all $n \in \mathbb{N}$. In other words, repeatability means that the reference's ability to perform the protocol never degrades, but crucially its state is allowed to change. Essentially this means that despite that the free energy of the reference can fluctuate and its coherence properties change, it can be used indefinitely to repeat the same protocol. To design such a protocol we employ the recent surprising result of [31] that shows how a coherence resource can be used repeatably to lift the symmetry constraints imposed by energy conservation⁷. However, as we will see the protocol in [31] requires continuous injection of energy into the reference (we do not allow the Hamiltonian of the reference to be unbounded from below, as in [47]). Hence, it is not immediately obvious that net thermodynamic work can be extracted from coherence.

In what follows we introduce a general protocol that processes quantum systems individually and allow us to extract work from their coherence. We then focus on two variations of it. The first one can come arbitrarily close to extracting all the coherence as average work with arbitrarily small failure probability, provided we make the coherence resources of the reference system in the thermal machine large enough. However, if one does not have access to arbitrarily large coherence resources, this variation of the protocol does not guarantee perfect repeatability. Therefore, we examine a second variation that is perfectly repeatable even for bounded references. We show then that the performance of work extraction in both the single-shot and asymptotic regimes is enhanced only if the quality of the reference (defined further in the text) is above a certain threshold.

3. The protocol

We analyse work extraction from pure qubit states with coherence

$$|\psi\rangle = \sqrt{1-p}|0\rangle + \sqrt{p}e^{-i\varphi}|1\rangle, \quad p \in (0, 1). \quad (6)$$

Without loss of generality we can set $H_S = |1\rangle\langle 1|$ and $\varphi = 0$ (rotations about the z axis of the Bloch sphere conserve energy). Our aim is to unlock work from coherence through the repeatable use of a thermal machine containing a reference, while processing each copy of $|\psi\rangle$ individually. In table 1 the results we obtain within this framework are schematically compared with the ones obtained within the frameworks presented in the previous section for the paradigmatic example of a qubit in a 'coherent Gibbs state' $|\gamma\rangle$ given by:

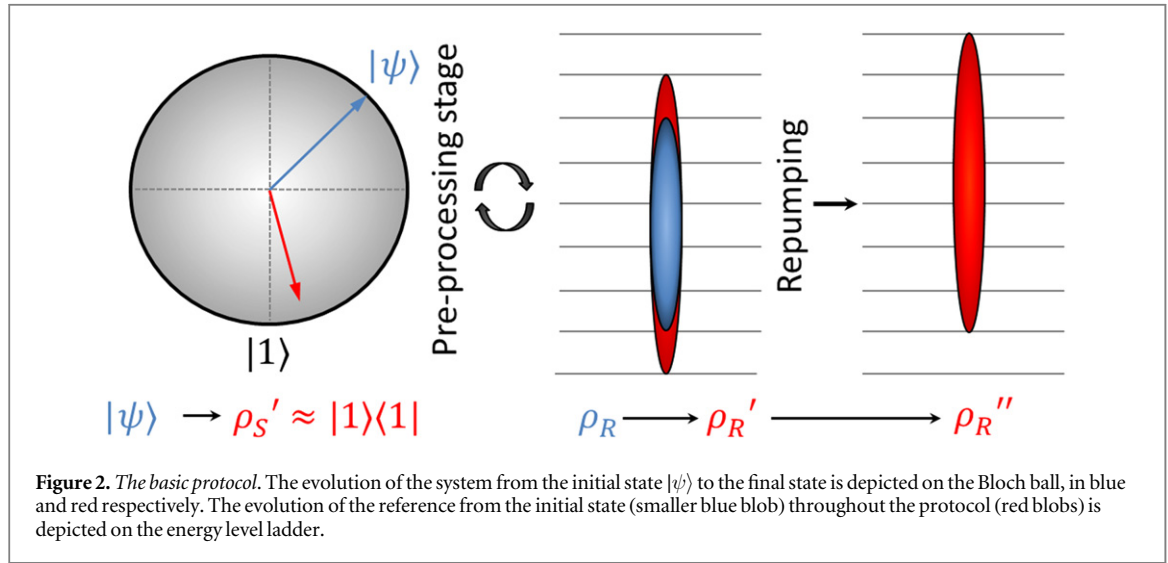
$$|\gamma\rangle = \sqrt{1-r}|0\rangle + \sqrt{r}|1\rangle, \quad (7)$$

with $(1-r, r)$ being the thermal distribution for the system, so that $\mathcal{D}(|\gamma\rangle\langle\gamma|) = \gamma_S$.

⁷ The work [31] actually uses the word 'catalysis', but we prefer to use the word repeatability/repeatable to avoid suggesting that there is no change in the state of the reference. Recall that traditionally a catalyst is a system in a state χ that enables $\rho \otimes \chi \rightarrow \sigma \otimes \chi$, despite $\rho \rightarrow \sigma$ being impossible (see, e.g., [8, 46]). Repeatability, on the other hand, only requires the auxiliary system to be as useful at the end as it was at the beginning, while its state may change.

Table 1. Individual processing protocols extracting work from $|\gamma\rangle$. $\langle W \rangle$ denotes the average work that can be extracted from the coherent thermal state $|\gamma\rangle$ and ϵ denotes the error probability of a single-shot work extraction from a given state. A thermal state of the system is denoted by γ_S . Note that under operations strictly conserving energy, no work extraction protocol on $|\gamma\rangle$ can outperform a work extraction protocol on γ_S , as the two states are indistinguishable.

	Single-shot	Asymptotic
Average energy conservation	$\langle W \rangle = \Delta F$ [35] Large fluctuations [9]	$\langle W \rangle = \Delta F$ [35]
Strict energy conservation	$\epsilon(\gamma\rangle) = \epsilon(\gamma_S), \langle W \rangle = 0$ [7, 10, 28]	
Strict energy conservation with resource used repeatably	$\epsilon(\gamma\rangle) = 0$ for unbounded reference $\epsilon(\gamma\rangle) < \epsilon(\gamma_S)$ for bounded reference	$\langle W \rangle = \Delta F$ for unbounded reference $\langle W \rangle < \Delta F$ for bounded reference



The extraction of non-zero work from $|\psi\rangle$ then requires a thermal machine containing a reference state ρ_R and implementing an energy-conserving unitary V :

$$\rho'_{SR} = V(|\psi\rangle\langle\psi| \otimes \rho_R)V^\dagger,$$

satisfying the following:

- (i) The system is pre-processed to a new state $\rho'_S = \text{Tr}_R(\rho'_{SR})$ that allows for better work extraction than from the initial state $|\psi\rangle$.
- (ii) The final reference state $\rho'_R = \text{Tr}_S(\rho'_{SR})$ can be processed into a state ρ''_R (perhaps using some of the extracted work) in such a way that the repeatability requirement is satisfied.
- (iii) No collective operations, at any stage of the protocol, are allowed on multiple copies of $|\psi\rangle$ and no quantum memory (in the sense of appendix A) is used.

3.1. The explicit work-extraction protocol

A protocol satisfying the introduced requirements consists of the following steps (see figure 2):

- (i) *Pre-processing.* The system $|\psi\rangle$ interacts through an energy-preserving unitary $V(U)$ with the reference ρ_R . The unitary acting on the joint system SR is chosen as in [31] to be:

$$V(U) = |0\rangle\langle 0| \otimes |0\rangle\langle 0| + \sum_{l=1}^{\infty} V_l(U), \quad (8)$$

with

$$V_l(U) = \sum_{n,m=0}^1 \langle n| U |m\rangle |n\rangle\langle m| \otimes |l-n\rangle\langle l-m|.$$

We choose

$$U = \begin{pmatrix} \sqrt{p} & -\sqrt{1-p} \\ \sqrt{1-p} & \sqrt{p} \end{pmatrix},$$

so that U rotates the qubit system from $|\psi\rangle$ to $|1\rangle$.⁸

- (ii) *Work extraction.* The system is now in a state ρ'_S and, due to work-locking (see section 2.2), is indistinguishable from its dephased version in any work extraction protocol. So without loss of generality we can use the dephased version

$$\mathcal{D}(\rho'_S) = (1-q)|0\rangle\langle 0| + q|1\rangle\langle 1|. \quad (9)$$

Now, depending on the considered regime, single-shot or average (asymptotic) work can be extracted from the dephased state and stored in the thermal machine.

- (iii) *Repumping.* The back-reaction changes the state of the reference into ρ'_R . Using part of the extracted work (stored in the battery during the previous step) we can repump the reference to shift it up:

$$\rho'_R \rightarrow \rho''_R := \Delta \rho'_R \Delta^\dagger, \quad (10)$$

with the details of how to perform such operation given in appendix C. We will describe how often we perform the repumping while analyzing different variations of the protocol.

- (iv) We can repeat the protocol using ρ''_R and a fresh copy of $|\psi\rangle$.

3.2. Performance

During the pre-processing stage the joint unitary $V(U)$ approximately induces U on the system:

$$\rho'_S \approx U(|\psi\rangle\langle\psi|)U^\dagger = |1\rangle\langle 1|.$$

The degree to which the above equation holds depends on the *quality* of the reference as defined in [31]. In particular, the system final occupation in the excited state $q = \langle 1| \rho'_S |1\rangle$ is given by

$$q = 1 - 2p(1-p)(1 - \langle \bar{\Delta} \rangle) - (1-p)^2 R_{00}, \quad (11)$$

where $R_{00} = \langle 0|\rho_R|0\rangle$ and $\bar{\Delta}$ is the quality parameter from definition 1. From equation (11) it is easy to see that $q \rightarrow 1$ when $R_{00} \rightarrow 0$ (i.e., when the reference quality parameter $M > 0$) and $\langle \bar{\Delta} \rangle \rightarrow 1$. Therefore, R_{00} and $\langle \bar{\Delta} \rangle$ are operationally well-defined quality parameters of the reference, because they directly measure the ability of the reference to induce the unitary U that we want to perform on the qubit. At the same time the reference undergoes a back-reaction induced by the joint unitary $V(U)$. This is described by the following Kraus operators:

$$A_0 = (1 - \sqrt{p})\sqrt{1-p} |0\rangle\langle 0| + \sqrt{p(1-p)}(\mathbb{I} - \Delta), \quad A_1 = p\mathbb{I} + (1-p)\Delta^\dagger, \quad (12)$$

so that the reference state after performing the pre-processing stage is given by:

$$\rho'_R = A_0 \rho_R A_0^\dagger + A_1 \rho_R A_1^\dagger. \quad (13)$$

From equations (9) and (11) the only two parameters relevant for work extraction are the reference population in the ground-state, R_{00} , and the parameter $\langle \bar{\Delta} \rangle$. Using the Kraus operators specified by equation (12), the change in $\langle \bar{\Delta} \rangle$ during the pre-processing stage (i.e., the difference between the final and initial value of $\langle \bar{\Delta} \rangle$) can be computed and is found to be:

$$\delta \langle \bar{\Delta} \rangle = (1-p)[(1 + \sqrt{p} - 2p)\text{Re}(R_{01}) - \sqrt{p}R_{00}], \quad (14)$$

where as before $R_{ij} = \langle i|\rho_R|j\rangle$. A sufficient condition for $\langle \bar{\Delta} \rangle$ to stay constant is $R_{00} = 0$, i.e., $M > 0$. Therefore, if the initial state satisfies $R_{00} = 0$ exactly, performing the pre-processing stage does not change $\langle \bar{\Delta} \rangle$, as noted in [31]. We require step (iii) of the protocol to ensure that $\langle 0|\rho''_R|0\rangle = 0$. If this is the case, at the end of the protocol we are left with the reference described by the same quality parameters R_{00} and $\langle \bar{\Delta} \rangle$ as at the beginning, and the reference ρ''_R is as good as ρ_R within the protocol.

⁸ This interaction corresponds to a modified Jaynes–Cummings model (with excitation-dependent coupling strengths). However, it can also be approximately realised within the standard Jaynes–Cummings model using a reference in a coherent state $|\alpha\rangle$, with $|\alpha|$ large enough (for details see supplementary material section V in [31]).

Finally, because the state of the reference changes, its free energy can fluctuate. However, notice firstly that the reference has Hamiltonian bounded from below, so for fixed average energy it has a finite amount of free energy. Secondly, repeatability requires that the reference can be used an arbitrary number of times and the performance of the protocol never changes. It is then easy to see that on average the free energy of the reference cannot be extracted as work, as this would be incompatible with repeatability. It can be shown that in the worst-case scenario the free energy change in the reference fluctuates around zero (see appendix D for more details). Moreover, these fluctuations vanish in the limit of an unbounded reference $\langle \bar{\Delta} \rangle \rightarrow 1$, as then the entropy of the reference stays constant, while its average energy increases. Therefore its free energy must increase at every step of the protocol.

4. Fundamental limitations of coherence to work conversion

How well does the above approach do in terms of work extraction? Here, we first emphasise the limitations on work extraction from coherence that arise due to the reference being bounded, i.e., when we have access to limited coherence resources. More precisely, we will explain why the use of a bounded reference does not allow one to extract from a state with coherence the average amount of work equal to the free energy difference.

However, we will then also show that one can construct a series of bounded reference states that come arbitrarily close to extract the free energy $\Delta F(|\psi\rangle)$, with protocols arbitrarily close to perfect repeatability⁹. Thus, we will prove that in the limit of an unbounded reference all coherence can be converted into work in a repeatable way. The limit case does not generate any entropy in the reference system and, being a reversible transformation, is optimal.

4.1. Limitations of bounded thermal machines

In order to illustrate the limitations arising from using a bounded reference we will consider a particular model of work extraction from coherence described in [35, 36]. It has been proved there that the free energy difference $\Delta F(\rho_S)$ can be extracted from a system ρ_S as work if one allows for the use of operations that conserve the energy only on average. Let us briefly recall the protocol used to achieve this. It is composed of two stages; first, given a state ρ_S , work is extracted from coherence. The resultant state is given by $\mathcal{D}(\rho_S)$ (recall that \mathcal{D} denotes the dephasing operation in the energy eigenbasis). Second, work is extracted from the incoherent state $\mathcal{D}(\rho_S)$. In accordance with assumption 2, the latter process extracts $\Delta F(\mathcal{D}(\rho_S))$. Hence, the extraction of the full free energy $\Delta F(\rho_S)$ from a state ρ_S is equivalent to the possibility of extracting $F(\rho_S) - F(\mathcal{D}(\rho_S))$ from the coherence of ρ_S . Notice that this quantity coincides with $kTA(\rho_S)$, where $A(\rho_S) = S(\rho_S|\mathcal{D}(\rho_S))$ is a known measure of quantum coherence [48], and it quantifies the amount of free energy stored in coherence [28]. Hence, the amount of work that needs to be extracted on average from the coherence of a quantum state to achieve the free energy extraction limit for arbitrary quantum states is

$$W_{\text{coh}}(\rho_S) = kTA(\rho_S). \quad (15)$$

Without loss of generality we can write any state ρ_S as $\sum_n p_n |\psi_n\rangle\langle\psi_n|$ with $p_{n+1} \leq p_n$. Let us also denote the Hamiltonian of the system by $H_S = \sum_n E_n |E_n\rangle\langle E_n|$. In the protocol that allows to extract work from coherence in [35], the system ρ_S interacts with a weight system in a gravitational field via the unitary

$$U_{\text{ave}} = \sum_n |E_n\rangle\langle\psi_n| \otimes \Gamma_{\varepsilon_n}, \quad (16)$$

where Γ_{ε_n} is the shift operator on the weight system that shifts it in energy by $\varepsilon_n = \langle\psi_n|H_S|\psi_n\rangle - E_n$.

As U_{ave} does not strictly conserve energy, by assumption 1 it is not a free thermodynamic operation. One can instead ask if it can be achieved by an energy-preserving unitary $V(U_{\text{ave}})$ on a larger system that exploits some ancillary system ρ_A (that includes a battery). In other words, we are looking for the energy-preserving unitary $V(U_{\text{ave}})$ such that

$$\mathcal{E}(\rho_S) := \text{Tr}_A(V(U_{\text{ave}})(\rho_S \otimes \rho_A)V(U_{\text{ave}})^\dagger) = U_{\text{ave}} \rho_S U_{\text{ave}}^\dagger. \quad (17)$$

It is easy to show that due to the imposed constraints, the ancillary system ρ_A must carry quantum coherence. In fact, if ρ_A were incoherent, then the left-hand side of equation (17) would be a time-translation covariant quantum map (meaning that $[\mathcal{E}, \mathcal{D}] = 0$) [33], whereas the right-hand side is not.

Now the crucial point is that equation (17) cannot hold exactly unless ρ_A contains an unbounded reference. If ρ_A is bounded, then the reduced evolution of ρ_S is not exactly unitary and not all the energy change can be identified with work. To prove this, one can compute the von Neumann entropy of both sides of equation (17)

⁹ A similar result appears in [31], however it was based on using a reference system described by a doubly infinite ladder Hamiltonian. This left open the question if this limit is achievable by a system with a physically realisable Hamiltonian.

and notice that the mutual information $I(\rho_{SA}) = 0$. Note, however, that the only way this is possible is if $V(U_{\text{ave}}) = V_1 \otimes V_2$, and further we would need $V_1 = U_{\text{ave}}$. But this is not possible, because from the fact that $V(U_{\text{ave}})$ is strictly energy-preserving we can prove that V_1 and V_2 must both be as well. Hence, the right-hand side of equation (17) cannot be a unitary if ρ_A contains only a bounded reference frame. In fact, equation (17) can only hold as a limit case of using a larger and larger coherence resource. In summary, assumptions 1–3 together with identification of work with energy change during unitary processes, imply that without an unbounded reference the work extraction protocol from [35, 36] cannot extract an amount of work equal to $\Delta F(\rho_S)$ from a state with coherence.

The strength of this point is that we do not even need to require the repeatability of the protocol using the same ancillary system. In fact, the argument is rather general. In order to extract all the free energy from a state ρ_S one needs to transform it into a thermal state. This cannot be achieved by only changing the energy spectrum of H_S , but also requires the rotation of the energy eigenbasis, so that the system is incoherent at the end of the transformation. This can be performed perfectly only with the aid of an unbounded reference frame, because it involves unitaries that do not strictly conserve energy¹⁰.

4.2. Extracting work arbitrarily close to the free energy difference

A key fact about the Carnot efficiency is that, despite being achieved only by ideal heat engines that do not actually exist in Nature, we can get arbitrarily close to it through a sequence of real engines. In a similar spirit, we now construct a sequence of bounded thermal machines getting arbitrarily close to the coherence to work conversion limit set by equation (15). The main result of this Section can be summarised in a non-technical way as follows:

Theorem 1. *There exists a sequence of bounded thermal machines approaching the ideal coherence to work conversion of equation (15) with arbitrarily high probability of success and with an arbitrarily small change in the quality parameters. The limit case is reversible.*

As an immediate consequence of the fact that the limit case is reversible we have:

Corollary 1. *Equation (15) provides the ultimate limit of coherence to average work conversion.*

In the remaining part of this section we give more details about the result above, first of all specifying the technical claim and then the main steps of the proof (the details of calculations can be found in appendix E). We consider a sequence of reference states ρ_R that approach a classical reference. Consider an arbitrary reference state ρ_R . We will describe it by two parameters $(\langle \bar{\Delta} \rangle, M)$ according to definition 1.

We will now show how to perform the protocol described in the previous section to extract from any pure state $|\psi\rangle$ an amount of work per copy arbitrarily close to the free energy difference $\Delta F(|\psi\rangle)$, while succeeding with arbitrarily high probability and changing the quality of the reference only by a negligible amount. For simplicity, define $f(x) = -x - kT h_2(x)$, where $h_2(\cdot)$ denotes the binary entropy. Theorem 1 can be now made technically precise as follows:

Theorem 1'. Let ρ_R be an arbitrary reference state described by $(\langle \bar{\Delta} \rangle, M)$. In the presence of a thermal bath at temperature T and if M is large enough, there exists a protocol individually extracting from M copies of $|\psi\rangle$ (given by equation (6)) an average amount of work $M \langle W \rangle$, with

$$\langle W \rangle \geq \Delta F(|\psi\rangle) - f(2p(1-p)(1 - \langle \bar{\Delta} \rangle)) - O(M^{-\frac{1}{3}}).$$

The probability of success p_{succ} of the protocol is

$$p_{\text{succ}} \gtrsim [1 - 2p(1-p)(1 - \langle \bar{\Delta} \rangle)]^M,$$

and it changes the quality parameters of the reference as follows:

$$\delta M = 0, \quad \delta \langle \bar{\Delta} \rangle \leq 2\sqrt{1 - p_{\text{succ}}}.$$

Before presenting the proof of this theorem, let us first comment on its scope. Note that the same result holds when a reference $(\langle \bar{\Delta} \rangle, M)$ is used a number of times $M' < M$, as long as $M' \gg 1$ (this will be clear from the proof). In the case in which $M' > M$, we can apply the theorem every M uses of the reference. The changes in the quality parameters will eventually sum up, but the theorem gives a bound on them. Also, it will be clear from the proof that the failure of the protocol implies a destruction of the coherence properties of the reference.

¹⁰ A useful point of view is also given by the theory of quantum reference frames and recovery maps [41, 49].

We now prove the Theorem by constructing an explicit variation of protocol introduced in section 3.1 and showing that it performs as stated in the theorem. Recall that by an energy conserving unitary we can rotate $|\psi\rangle$ around the z axis of the Bloch sphere. Hence, without loss of generality, we can set $\varphi = 0$ in equation (6). We then perform steps (i) and (ii) of the protocol described in section 3.1 M times, i.e., individually processing each of M copies of $|\psi\rangle$ using a reference ρ_R described by $(\langle\bar{\Delta}\rangle, M)$. The choice of ρ_R ensures that during this process the reference state will have no population in the ground state, and so $\langle\bar{\Delta}\rangle$ will stay constant. Then, the final state of the reference is described with probability p_1^M by $\rho_{R,1} = A_1^M \rho_R A_1^{\dagger M}$, where $p_1 = (1 - 2p(1 - p)(1 - \langle\bar{\Delta}\rangle))$. Notice that having access to a reference described by parameters $\langle\bar{\Delta}\rangle$ and M such that $M(1 - \langle\bar{\Delta}\rangle) \rightarrow 0$, the probability p_1^M can be made arbitrarily close to 1. This happens because by taking $\langle\bar{\Delta}\rangle$ close enough to 1, we can get arbitrarily close to unitary evolution of a system state $|\psi\rangle$ to a pure, incoherent state $|1\rangle$. As the final state of the system is almost pure, the final joint state of the system and the reference factorises and an arbitrarily small amount of entropy is generated (the back-reaction on the reference is given by the Kraus operator A_1 alone).

Next, we repump the reference $\bar{M} + s\sigma_M^{4/3}$ times, where $\bar{M} = M(1 - p)$, $\sigma_M = \sqrt{Mp(1 - p)}$ and $s > 0$. This guarantees that the reference has arbitrarily small population in states $\{|0\rangle \dots |M\rangle\}$, so that by performing a measurement we can project the reference to a state ρ_R'' with support on the subspace spanned by $\{|i\rangle\}_{i>M}$ with arbitrarily high probability (s fixes the confidence level, see appendix E for details). More precisely, after repeating steps (i) and (ii) of the protocol M times and repumping as explained above, the reference is described by a state ρ_R'' with probability

$$p_{\text{succ}} \geq [1 - 2p(1 - p)(1 - \langle\bar{\Delta}\rangle)]^M \mathbb{E}_s(M^{1/6}), \quad (18)$$

where $\mathbb{E}_s(x) = \text{erf}(sx/\sqrt{2})$ and erf denotes the error function. The final state is given by ρ_R'' , described by $(\langle\bar{\Delta}\rangle'', M'')$, where $M'' = M$ and

$$\delta\langle\bar{\Delta}\rangle := \langle\bar{\Delta}\rangle'' - \langle\bar{\Delta}\rangle \leq 2\sqrt{1 - p_{\text{succ}}} \leq 2\sqrt{1 - p_1^M \mathbb{E}_s(M^{1/6})}. \quad (19)$$

Notice that by taking s large enough (but finite) we can make the factor $\mathbb{E}_s(M^{1/6})$ in the previous two equations arbitrarily close to 1, $\mathbb{E}_s(M^{1/6}) \approx 1$. In the appropriately chosen limit $\langle\bar{\Delta}\rangle \rightarrow 1$ and $M \rightarrow \infty$ the quality parameters of the reference state are then unchanged with probability 1. Let us also note that the cost W_E of the measurement described above is bounded by $kTh_2(p_{\text{succ}})$ (see appendix E).

We have just shown that following the procedure above we can guarantee repeatability with arbitrary confidence level. Hence, we now proceed to proving that it also allows for extracting an average amount of work per system arbitrarily close to the free energy difference $\Delta F(|\psi\rangle)$. To see this, note that after repeating the protocol on M copies of $|\psi\rangle$ we are left with M copies of a state $\mathcal{D}(\rho_S')$ from equation (9) with q given by equation (11), $q = 1 - 2p(1 - p)(1 - \langle\bar{\Delta}\rangle)$. This state is diagonal in the energy eigenbasis and the average work $\langle\tilde{W}\rangle$ extracted from it is given by $\Delta F(\mathcal{D}(\rho_S'))$:

$$\langle\tilde{W}\rangle = 1 + kT \log Z - 2p(1 - p)(1 - \langle\bar{\Delta}\rangle) - kTh_2(2p(1 - p)(1 - \langle\bar{\Delta}\rangle)), \quad (20)$$

where as before $h_2(\cdot)$ denotes the binary entropy. By choosing M large enough, we can ensure that the extracted work is arbitrarily peaked around the average given by the above equation (M can be bounded using the results of [9]). This ensures that when we need to repump the reference, we actually have enough work to invest to do it. The repumping costs $\bar{M} + s\sigma_M^{4/3}$ units of extracted work and the cost W_E of measurement is independent from M . Hence, the net gain per processed copy of $|\psi\rangle$ is given by

$$\langle\tilde{W}\rangle - \frac{(\bar{M} + s\sigma_M^{4/3}) + W_E}{M} = \langle W \rangle - O(M^{-1/3}), \quad (21)$$

where

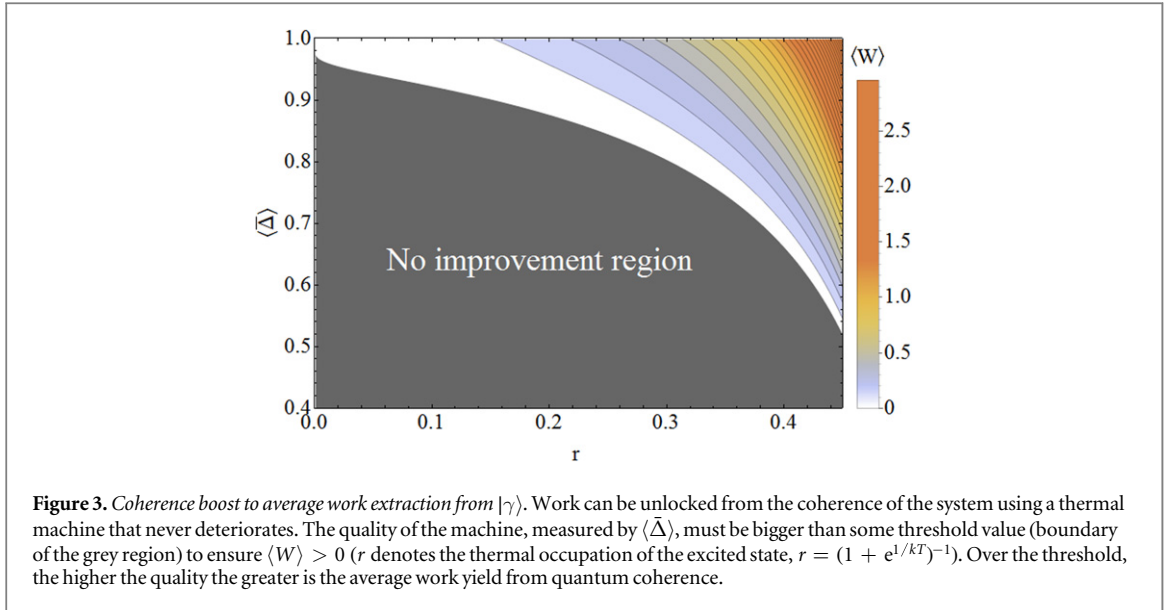
$$\langle W \rangle = \Delta F(|\psi\rangle) - 2p(1 - p)(1 - \langle\bar{\Delta}\rangle) - kTh_2(2p(1 - p)(1 - \langle\bar{\Delta}\rangle)). \quad (22)$$

Therefore, the deficit per copy scales as $M^{-1/3}$ and by choosing M large enough it can be made arbitrarily small. Moreover, the previous equation gives us the relation between the quality of the reference and the average extracted work, showing that $\langle W \rangle \rightarrow \Delta F(|\psi\rangle)$ as $\langle\bar{\Delta}\rangle \rightarrow 1$, $M \rightarrow \infty$, $M(1 - \langle\bar{\Delta}\rangle) \rightarrow 0$.

We conclude that it is possible, with arbitrarily large success probability, to extract an amount of work arbitrarily close to the free energy change from a pure state with coherence in energy eigenbasis, while processing it individually and properly taking account of all the resources used, i.e., ensuring arbitrarily exact repeatability.

5. Extracting work with perfect repeatability and bounded thermal machines

In the previous section we have shown how to extract as work all free energy of a pure quantum state with coherence. However we allowed for



- (i) The limit case of an unbounded thermal machine, $\langle \bar{\Delta} \rangle \rightarrow 1$.
- (ii) An asymptotic protocol individually processing a large number M of copies of the system.

These assumptions may be too strong if the reference itself is a microscopic system involved in the thermodynamic processing and exclude the applicability to single-shot scenarios. What if we only want to process a small number of systems? This requires us to go beyond the results of the previous section.

Moreover, even if we only want to release the first of the two assumptions, i.e., put a bound on the coherence properties of the reference, we are still left with open questions. In this case the general result stated by theorem 1 is applicable, however the work extraction protocols presented always entail a failure probability $1 - p_{\text{succ}}$ that can lead to a complete destruction of the coherent properties of the reference. Even if this probability is relatively small, we may not be willing to take this risk. Also, the reference inevitably deteriorates, even if by a small amount bounded by equation (19). A crucial question is then: Are there work extraction protocols with $\langle \bar{\Delta} \rangle < 1$ such that $\delta \langle \bar{\Delta} \rangle = 0$ and $p_{\text{succ}} = 1$? In other words, can we extract work from coherence using a protocol that *never* fails and gives back the thermal machine with *exactly* the same quality parameters, even if the reference is bounded?

In this section we construct such protocols for both average and single-shot work extraction. These ensure perfect repeatability, but the price we pay is that the average amount of extracted work is strictly smaller than the free energy difference and it is only possible for $\langle \bar{\Delta} \rangle$ above a certain threshold value $\bar{\Delta}_{\text{crit}}$. In the case of single-shot work extraction we show similarly that there exists a threshold over which the reference allows us to outperform the single-shot protocol with no coherence. For clarity of the discussion, we focus on the paradigmatic case of the class of states $|\gamma\rangle$ introduced in equation (7).

5.1. Average work extraction

In absence of an external source of coherence no work can be extracted from the state $|\gamma\rangle$ on average [10, 28]. However, if we allow for a repeatable use of the reference, positive work yield can be obtained. In order to achieve this, during step (ii) of the protocol we perform average work extraction from the state $\mathcal{D}(\rho'_S)$ specified by equation (9). As $\mathcal{D}(\rho'_S)$ is diagonal in the energy eigenbasis, the results of [9, 10] apply. Therefore, the average work yield is given by the free energy difference $\Delta F(\mathcal{D}(\rho'_S))$. To ensure perfect repeatability we repump the reference at each run, so that if $\langle 0|\rho_R|0\rangle = 0$, then $\langle 0|\rho'_R|0\rangle = 0$, and the reference quality parameters do not change. The repumping requires a unit of work, so that the work extracted on average during one run of the protocol is

$$\langle W \rangle = q + kT (\log Z_S - h_2(q)) - 1. \quad (23)$$

The connection between the properties of the reference and the work yield is given by equation (23) together with equation (11) (where $R_{00} = 0$ and $p = r$).

In figure 3 we show how much work $\langle W \rangle$ can be unlocked through our protocol as a function of the quality of the reference $\langle \bar{\Delta} \rangle$ and the thermal occupation r of the excited state. The graph shows that the quality of the reference needs to be above a certain threshold in order to get positive average work yield. As expected, the

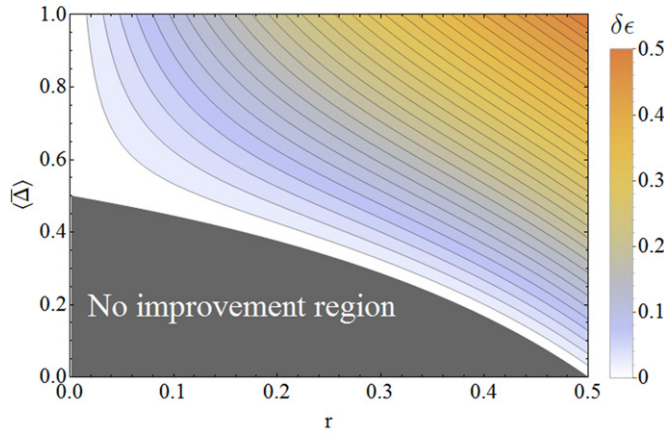


Figure 4. Coherence boost to single-shot work extraction from $|\gamma\rangle$. A thermal machine, used in a repeatable way, can exploit the quantum coherence of the system to decrease the failure probability in single-shot work extraction. Similarly to the case of average work extraction, the quality of the machine $\langle\bar{\Delta}\rangle$ must be over some threshold value to lead to any improvement (r denotes the thermal occupation of an excited state). As $\langle\bar{\Delta}\rangle$ increases, the failure probability decreases from r to $r - \delta\epsilon$, down to zero, i.e., to the point when the single-shot work extraction from the pure quantum state $|\gamma\rangle$ becomes fully deterministic.

advantage is the most significant for high r , because the states $|\gamma\rangle$ and γ_S differ most in this case or, in other words, the amount of coherence to be unlocked is higher.

As already mentioned in section 2, in the asymptotic regime of individually processing large number of copies of $|\gamma\rangle$, the fluctuations in the work yield, equation (23), become negligible. Notice, however, that even if $\langle W \rangle > 0$ we may not be able to perform step (iii) every time, as the fluctuations around the average mean that we will not always have enough work to invest in the repumping. To resolve this problem we can follow a strategy analogous to the case of unbounded reference. That is, we repump after having extracted work M times, where M is sufficiently large to neglect the fluctuations around $\langle W \rangle$.¹¹ The protocol will be repeatable up to an arbitrarily small probability of failure, if the support of the reference initially starts high enough in the energy ladder. It is important to stress, however, that the ‘failure’ in this case does not entail a destruction of the coherence properties of the reference, as in section 4.2. It only requires the investment of extra work in order to ensure perfect repeatability.

5.2. Single-shot work extraction

Finally, we proceed to a fully quantum and single-shot protocol for an individual quantum state. This version of the protocol does not assume possessing an unbounded reference nor it requires asymptotic number of runs. In absence of an external source of coherence we can perform ϵ -deterministic work extraction from $|\gamma\rangle$ – as $|\gamma\rangle$ is indistinguishable from γ_S , the results of [9] apply. This means that we can extract $kT \log Z_S$ work with failure probability r or $1 + kT \log Z_S$ with failure probability $1 - r$. We now show that exploiting the reference in a perfectly repeatable way the failure probability for extracting $kT \log Z_S$ can be decreased—and the higher the quality of the reference, the stronger the improvement.

During step (ii) of the protocol we perform ϵ -deterministic work extraction from the state $\mathcal{D}(\rho'_S)$ specified by equation (9), in accordance with assumption 3. With probability q we extract $1 + kT \log Z_S$ work and with probability $1 - q$ our protocol fails. As we need one unit of work to repump the reference (see equation (10)), the net gain is $kT \log Z_S$. When the protocol fails (with probability q), the reference is returned in the state ρ'_R and one has to invest one unit of work to ensure repeatability.

In figure 4 we present the decrease $\delta\epsilon$ in the failure probability ϵ achieved by our protocol as compared to work extraction from γ_S . We see that if the quality of the reference is high enough, the coherence content of $|\gamma\rangle$ can be exploited to provide an advantage in the work extraction. In the limit of a very high quality (unbounded) reference, $\langle\bar{\Delta}\rangle \rightarrow 1$, the failure probability can be sent to zero, i.e., the work extraction from $|\gamma\rangle$ becomes deterministic.

¹¹ One can think of alternative protocols as well, in which at every repetition we toss a coin to decide if we repump the reference or not. We do not delve into this, but we expect to find similar results.

6. Conclusions

In this work we have addressed the following question: How much work can be extracted from a state that is a superposition of energy eigenstates? We argued that this question, within the currently developed theory of thermodynamics of individual quantum systems, is a subtle issue. We showed that the optimal coherence to work conversion can be obtained only in the limit of accessing a reference system with unbounded coherence resources. Although no real reference is unbounded (in the same way in which no heat engines is ideal), we can get arbitrarily close to the limit by means of a sequence of bounded thermal machines.

The access to arbitrarily large resources should be questioned in the regime under study. Generally speaking, recovering traditional thermodynamical results requires extra assumptions, which all entail some notion of ‘classicality’, so to effectively make coherence negligible: neglecting the energy fluctuations due to superposition; assuming the existence of a source of coherence that experiences essentially no back-reaction; collectively operating on infinitely many copies of the system.

When the ‘classicality’ assumptions are dropped one after the other, the results are quantitatively different from the thermodynamics of incoherent systems. Nevertheless, we find that the coherence between energy levels can still enhance the performance of work extraction protocols. There exist perfectly repeatable processes extracting on average a larger amount of work that could be extracted in the absence of coherence; and single-shot protocols in which coherence improves the success probability of work extraction. Although these protocols are better than the correspondent incoherent ones, they do not achieve the performance reached in the classical limit.

We also point out that while dealing with microscopic systems, the accounting of all the resources involved in thermodynamic processes becomes a crucial and non-trivial task. In this regard, we underline the importance of accounting for the resources which make up a thermal machine, and the concept of repeatability, that essentially captures the idea of using these extra resources without degrading them. In particular, the considerations here suggest that a full theory of thermodynamics in the quantum regime will require a better understanding of the accounting of coherence resources, including those found in the thermal machine. Some laws which place restrictions on coherence have been introduced in [8, 28, 30, 32, 33], but we are still far from having a full understanding. We hope to have convinced the reader that the question of the role of quantum coherence in thermodynamic considerations does not admit an easy and immediate answer, and that it is only by appropriately incorporating it into the theoretical framework that we can explore truly quantum mechanical effects.

Acknowledgments

We thank Terry Rudolph, Tony Short and Philipp Kammerlander for helpful discussions. We would also like to thank Janet Anders for her useful comments on a draft of this paper. Finally, we are grateful for constructive criticism and suggested changes provided by an anonymous referee. KK and ML are supported by EPSRC and in part by COST Action MP1209. DJ is supported by the Royal Society. JO is supported by an EPSRC Established Career Fellowship.

Appendix A. Collective processing regime

In the collective processing regime work can be effectively unlocked from coherence. This is achieved by processing many copies of a system state ρ_S collectively and extracting work from relational degrees of freedom that live in decoherence-free subspaces [10, 26, 28, 41]. The intuitive explanation is that one copy of a state ρ_S with coherence can act as a reference for the other one, and we have $\mathcal{D}(\rho_S^{\otimes 2}) \neq \mathcal{D}(\rho_S)^{\otimes 2}$. In the case of finite number of copies $\rho_S^{\otimes N}$ a non-zero amount of work is unlocked from the coherences, and in the limit of processing collectively infinitely many independent and identically distributed (i.i.d.) copies, the amount of work per copy that can be extracted deterministically equals $F(\rho_S) - F(\gamma_S)$.

Instead of collectively processing many copies of a system, one may also consider a black-box device B that takes in individual quantum systems ρ_S and at *each round* returns a thermalised state γ_S and an average amount of work equal to $F(\rho_S) - F(\gamma_S)$. From outside the box it seems we are dealing with a work extraction protocol that individually processes each state. However, the devious way in which the box achieves this is the following:

- (i) The box B contains a large quantum memory consisting of $N \gg 1$ copies of incoherent quantum states $\sigma_B^{\otimes N}$, for which $F(\sigma_B) = F(\rho_S)$.
- (ii) Every time the box takes in a single copy ρ_S it swaps this state into memory and instead performs work extraction on one copy of σ_B . Hence it outputs on average $F(\rho_S) - F(\gamma_S)$ and a thermalised state γ_S .

- (iii) After N uses its memory is filled with the coherent states $\rho_B^{\otimes N}$ and so it does large- N collective processing and restores to $\sigma_B^{\otimes N}$ with costs growing only sublinearly with N [26].

Although from the outside of the box this is identical to the individual processing regime, the collective, relational processing of coherence is ‘hidden’ in the quantum memory.

Appendix B. Average energy conservation does not explicitly model energy fluctuations

Consider a system and an ancilla described by Hamiltonians H_S and H_A , and prepared in states ρ_S and ρ_A , respectively. Assume also that the initial state of the system ρ_S has coherence between energy eigenspaces. Now consider a joint energy-conserving unitary U , i.e., $[U, H_S + H_A] = 0$, inducing the following evolution:

$$U(\rho_S \otimes \rho_A)U^\dagger = \rho'_S \otimes \rho'_A,$$

so that the final state of the system ρ'_S has no coherence in the energy eigenbasis and $\rho'_S = V\rho_S V^\dagger$ for some unitary V that conserves average energy. The uncertainty of an energy measurement on ρ'_A can be decomposed as [50]:

$$H(\rho'_A) = S(\rho'_A) + A(\rho'_A), \quad (24)$$

where $A(\sigma) = S(\sigma || \mathcal{D}(\sigma))$ is the relative entropy between a state and its decohered version and H is the Shannon entropy of the probability distribution of an energy measurement. Because U commutes with the total Hamiltonian we have

$$A(\rho_S \otimes \rho_A) = A(\rho'_S \otimes \rho'_A).$$

As the final state of the system ρ'_S has no coherence we have $A(\rho'_S \otimes \rho'_A) = A(\rho'_A)$. Using $A(\rho_S \otimes \rho_A) > A(\rho_A)$, one gets that $A(\rho'_A) > A(\rho_A)$. From the invariance of the von Neumann entropy under unitary transformations, $S(\rho'_A) = S(\rho_A)$. So we conclude from equation (24)

$$H(\rho'_A) > S(\rho_A) + A(\rho_A) = H(\rho_A).$$

Appendix C. Details of the repumping stage

Although one could question the repumping stage described by equation (10), given that Δ is not a unitary, we note that this is actually not a problem. This is because such operation can be realised through a joint energy-conserving unitary between a weight system in a state $|1\rangle$ and the reference in a state ρ'_R . The unitary is given by $V(U)$ in equation (8), where we take $U = X$, the Pauli X operator. Then

$$V(X) = \sigma_- \otimes \Delta + \sigma_+ \otimes \Delta^\dagger + |0\rangle\langle 0| \otimes |0\rangle\langle 0|,$$

where $\sigma_+ = |1\rangle\langle 0|$ and $\sigma_- = |0\rangle\langle 1|$. As the reference has no population in the ground state, the final state of the weight system is $|0\rangle$ and the final state of the reference is given by equation (10).

Appendix D. Free energy change of the reference

Denote by $\Delta F_{R,n}$ the change in the free energy of the reference at the n th repetition of the protocol. The total free energy change of the reference after M repetitions of the protocol satisfies

$$\sum_{n=1}^M \Delta F_{R,n} \leq F(\rho_R) - F(\gamma_R) \quad \forall M,$$

where γ_R is the thermal state of the reference. Hence, the average change in the free energy of the reference as $M \rightarrow \infty$ is

$$\overline{\Delta F_R} := \lim_{M \rightarrow \infty} \frac{1}{M} \sum_{n=1}^M \Delta F_{R,n} = 0.$$

Appendix E. Details of approaching free energy limit

We provide here the details of the repumping protocol. We start from a generic reference state ρ_R such that

$$\text{supp}(\rho_R) \cap \text{span}\{|0\rangle, \dots, |M\rangle\} = \emptyset,$$

where $|i\rangle$ are eigenstates of the reference Hamiltonian. We impose the requirements $Mp \gg 1$ and $M(1-p) \gg 1$, where p is fixed by equation (6).

We now compute the probability of the occurrence of the Kraus A_1 on a generic reference state σ . From equation (12) and the fact that $\Delta\Delta^\dagger = \mathbb{I}$ we obtain

$$p_1(\sigma) := \text{Tr}(A_1\sigma A_1^\dagger) = 1 - 2p(1-p)(1 - \text{Tr}(\bar{\Delta}\sigma)),$$

where recall that $\bar{\Delta} = (\Delta + \Delta^\dagger)/2$. Define the state of the reference after performing work extraction on $n \geq 1$ qubits through the following recurrence formula

$$\rho_R^{(n)} := A_0\rho_R^{(n-1)}A_0^\dagger + A_1\rho_R^{(n-1)}A_1^\dagger, \quad (25)$$

where $\rho_R^{(0)} = \rho_R$. Because ρ_R has initially no support in the first M energy levels, we can extract work from M qubits before there is any overlap with the ground state. In other words, $\langle 0|\rho_R^{(n)}|0\rangle = 0$, $\forall n \in \{1, \dots, M\}$. The previous formula, together with equation (14), implies that $\bar{\Delta}$ is conserved throughout the protocol. Hence we deduce that

$$\text{Tr}(A_1^M \rho_R A_1^{\dagger M}) = p_1^M(\rho_R) = (1 - 2p(1-p)(1 - \langle \bar{\Delta} \rangle))^M.$$

For notational convenience, we will now drop the explicit dependence of p_1 on ρ_R (initial state of the reference).

Using equation (25) we have

$$\rho_R^{(M)} = p_1^M \rho_{R,1} + (1 - p_1^M) \mathcal{E}_{\text{else}}^{(M)}(\rho_R), \quad (26)$$

where $\mathcal{E}_{\text{else}}^{(M)}$ contains all strings of A_0 's and A_1 's different from the string consisting only of A_1 's and

$$\rho_{R,1} = A_1^M \rho_R A_1^{\dagger M} / p_1^M.$$

We can now compute A_1^M :

$$A_1^M = \sum_{k=0}^M \binom{M}{k} p^{M-k} (1-p)^k \Delta^{\dagger k}.$$

We see that A_1^M is binomially distributed in the number of lowering operations Δ^\dagger . The average number of lowerings is $\bar{M} = M(1-p)$ and the standard deviation is $\sigma_M = \sqrt{Mp(1-p)}$. We can perform a number of repumpings as in equation (10) as detailed in appendix C. Let us denote this operation by \mathcal{P} . We have chosen M sufficiently large so that the confidence levels associated to σ_M are approximately gaussian. Hence, we can repump the reference $\bar{M} + s\sigma_M^{4/3}$ times, which guarantees that the reference has arbitrarily small population in states $|0\rangle \dots |M\rangle$ with a confidence level controlled by $s > 0$ and increasing with M . More precisely, if P_M is the projector on the subspace spanned by $\{|0\rangle, \dots, |M\rangle\}$ and $P_M^\perp = \mathbb{I} - P_M$,

$$\text{Tr}(P_M^\perp \mathcal{P}(\rho_{R,1})) \geq \text{erf}(sM^{1/6}/\sqrt{2}) := \mathbb{E}_s(M^{1/6}), \quad (27)$$

where $\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2) dt$ denotes the error function. Now, using equations (26) and (27)

$$p_{\text{succ}} := \text{Tr}(P_M^\perp \mathcal{P}(\rho_R^{(M)})) \geq p_1^M \mathbb{E}_s(M^{1/6}). \quad (28)$$

This implies that in performing the two-outcome measurement $\{P_M, P_M^\perp\}$ we would find the outcome P_M^\perp with probability given by equation (28).

Performing such a measurement guarantees that the final state of the reference will have no support on a subspace spanned by $\{|0\rangle \dots |M\rangle\}$, similarly to the initial state. However, performing a selective measurement has a thermodynamic cost that we have to take into account. More precisely, such a measurement can be performed using an ancillary memory qubit system A described by trivial Hamiltonian $H_A = 0$. Then, taking the initial state of A to be a pure state $|0\rangle$, we can perform operation on the joint reference-ancillary state described by the Kraus operators $M_1 = P_M^\perp \otimes \mathbb{I}$ and $M_2 = P_M \otimes \sigma_x$. This operation is energy conserving, as the Kraus operators commute with the total Hamiltonian $H_R + H_A$. Hence, it is free of thermodynamic cost. Now the projective measurement on states $|0\rangle$ and $|1\rangle$ can be performed on the ancillary memory system. Observing the result 0 will project the reference on a subspace P_M^\perp , whereas observing the result 1 will project the reference on P_M . The thermodynamic cost associated with this projective measurement is the cost of erasing the memory system afterwards. This is given by

$$W_E = kT h_2(p_{\text{succ}}),$$

which can be made arbitrarily small as $p_{\text{succ}} \rightarrow 1$. Notice that we only needed to use a classical memory to record the measurement outcome, which is not in contrast with assumption (iii) of section 3. Also note that this cost has to be paid only after extracting work from M copies, hence the cost per copy scales as M^{-1} .

Define

$$\rho_R'' := \frac{P_M^\perp \mathcal{P}(\rho_R^{(M)}) P_M^\perp}{\|P_M^\perp \mathcal{P}(\rho_R^{(M)}) P_M^\perp\|}.$$

Now, using the gentle measurement lemma [51, 52], equation (28) also implies

$$\|\rho_R'' - \mathcal{P}(\rho_R^{(M)})\| \leq 2\sqrt{1 - p_{\text{succ}}}. \quad (29)$$

From equation (29), and the following characterisation of the trace norm (see [53])

$$\|\rho - \sigma\| = \max_{0 \leq A \leq \mathbb{I}} \text{Tr}(A(\rho - \sigma))$$

we find that

$$\text{Tr}(\bar{\Delta} \rho_R'') \geq \text{Tr}(\bar{\Delta} \mathcal{P}(\rho_R^{(M)})) - 2\sqrt{1 - p_{\text{succ}}} = \text{Tr}(\bar{\Delta} \rho_R) - 2\sqrt{1 - p_{\text{succ}}},$$

where the last equality comes from the fact that $\langle \bar{\Delta} \rangle$ is conserved in the protocol, up to the measurement. The last equation can be rewritten as

$$\text{Tr}(\bar{\Delta}(\rho_R - \rho_R'')) \leq 2\sqrt{1 - p_{\text{succ}}}.$$

Exchanging the roles of ρ_R and ρ_R'' and introducing

$$\delta \langle \bar{\Delta} \rangle = \text{Tr}(\bar{\Delta} \rho_R'') - \text{Tr}(\bar{\Delta} \rho_R),$$

we conclude

$$|\delta \langle \bar{\Delta} \rangle| \leq 2\sqrt{1 - p_{\text{succ}}}.$$

Using equation (28), this bounds the maximum allowed change of the quality parameter of the reference.

References

- [1] Maxwell J C 1873 *Theory of Heat* (London: Longmans Green)
- [2] Szilard L 1929 *Z. Phys.* **53** 840–56
- [3] Janzing D, Wocjan P, Zeier R, Geiss R and Beth T 2000 *Int. J. Theor. Phys.* **39** 2717–53
- [4] Horodecki M, Horodecki P and Oppenheim J 2003 *Phys. Rev. A* **67** 062104
- [5] Dahlsten O, Renner R, Rieper E and Vedral V 2011 *New J. Phys.* **13** 053015
- [6] Del Rio L, Åberg J, Renner R, Dahlsten O and Vedral V 2011 *Nature* **474** 61–3
- [7] Horodecki M and Oppenheim J 2013 *Nat. Commun.* **4** 2059
- [8] Brandão F G S L, Horodecki M, Ng N H Y, Oppenheim J and Wehner S 2015 *Proc. Natl Acad. Sci. USA* **112** 3275
- [9] Åberg J 2013 *Nat. Commun.* **4** 1925
- [10] Skrzypczyk P, Short A J and Popescu S 2013 arXiv:1302.2811
- [11] Halpern N Y, Garner A J, Dahlsten O C and Vedral V 2015 *New J. Phys.* **17** 095003
- [12] Egloff D, Dahlsten O C O, Renner R and Vedral V 2015 *New J. Phys.* **17** 073001
- [13] Faist P, Dupuis F, Oppenheim J and Renner R 2015 *Nat. Commun.* **6** 7669
- [14] Lostaglio M, Müller M P and Pastena M 2015 *Phys. Rev. Lett.* **115** 150402
- [15] Renes J M 2014 *Eur. Phys. J. Plus* **129** 153
- [16] Gemmer J and Anders J 2015 *New J. Phys.* **17** 085006
- [17] Collin D, Ritort F, Jarzynski C, Smith S B, Tinoco I and Bustamante C 2005 *Nature* **437** 231–4
- [18] Serreli V, Lee C, Kay E and Leigh D 2007 *Nature* **445** 523–7
- [19] Toyabe S, Sagawa T, Ueda M, Muneyuki E and Sano M 2010 *Nat. Phys.* **6** 988–92
- [20] Alemany A and Ritort F 2010 *Europhys. News* **41** 27–30
- [21] Cheng J, Sreelatha S, Hou R, Efremov A, Liu R, van der Maarel J R C and Wang Z 2012 *Phys. Rev. Lett.* **109** 238104
- [22] Lloyd S 2011 *J. Phys.: Conf. Ser.* **302** 012037
- [23] Lambert N, Chen Y N, Cheng Y C, Li C M, Chen G Y and Nori F 2013 *Nat. Phys.* **9** 10–8
- [24] Gauger E M, Rieper E, Morton J J, Benjamin S C and Vedral V 2011 *Phys. Rev. Lett.* **106** 040503
- [25] Scully M O, Zubairy M S, Agarwal G S and Walther H 2003 *Science* **299** 862–4
- [26] Brandão F G S L, Horodecki M, Oppenheim J, Renes J M and Spekkens R W 2013 *Phys. Rev. Lett.* **111** 250404
- [27] Rodríguez-Rosario C A, Frauenheim T and Aspuru-Guzik A 2013 arXiv:1308.1245
- [28] Lostaglio M, Jennings D and Rudolph T 2015 *Nat. Commun.* **6** 6383
- [29] Gelbwaser-Klimovsky D and Kurizki G 2014 *Phys. Rev. E* **90** 022102
- [30] Ćwikliński P, Studziński M, Horodecki M and Oppenheim J 2015 *Phys. Rev. Lett.* **115** 210403
- [31] Åberg J 2014 *Phys. Rev. Lett.* **113** 150402
- [32] Narasimhachar V and Gour G 2015 *Nat. Commun.* **6** 7689
- [33] Lostaglio M, Korzekwa K, Jennings D and Rudolph T 2015 *Phys. Rev. X* **5** 021001
- [34] Uzdin R, Levy A and Kosloff R 2015 *Phys. Rev. X* **5** 031044
- [35] Skrzypczyk P, Short A J and Popescu S 2014 *Nat. Commun.* **5** 4185
- [36] Kammerlander P and Anders J 2015 arXiv:1502.02673

- [37] Masanes L and Oppenheim J 2014 arXiv:[1412.3828](#)
- [38] Frenzel M F, Jennings D and Rudolph T 2014 *Phys. Rev. E* **90** 052136
- [39] Gallego R, Eisert J and Wilming H 2015 arXiv:[1504.05056](#)
- [40] Aharonov Y and Susskind L 1967 *Phys. Rev.* **155** 1428–31
- [41] Bartlett S D, Rudolph T and Spekkens R W 2007 *Rev. Mod. Phys.* **79** 555
- [42] Gelbwaser-Klimovsky D, Alicki R and Kurizki G 2013 *Europhys. Lett.* **103** 60005
- [43] Bartlett S D, Rudolph T, Spekkens R W and Turner P S 2006 *New J. Phys.* **8** 58
- [44] White G A, Vaccaro J A and Wiseman H M 2009 *AIP Conf. Proc.* **1110** 79–82
- [45] Ahmadi M, Jennings D and Rudolph T 2010 *Phys. Rev. A* **82** 032320
- [46] Horodecki R, Horodecki P, Horodecki M and Horodecki K 2009 *Rev. Mod. Phys.* **81** 865–942
- [47] Malabarba A S, Short A J and Kammerlander P 2015 *New J. Phys.* **17** 045027
- [48] Gour G, Marvian I and Spekkens R W 2009 *Phys. Rev. A* **80** 012307
- [49] Bartlett S D, Rudolph T, Spekkens R W and Turner P S 2009 *New J. Phys.* **11** 063013
- [50] Korzekwa K, Lostaglio M, Jennings D and Rudolph T 2014 *Phys. Rev. A* **89** 042122
- [51] Winter A 1999 *IEEE Trans. Inf. Theory* **45** 2481–5
- [52] Ogawa T and Nagaoka H 2002 arXiv:[quant-ph/0208139](#)
- [53] Nielsen M A and Chuang I L 2010 *Quantum Computation and Quantum Information* (Cambridge: Cambridge University Press)